

The Development of Early Academic Success: The Impact of Direct Instruction's *Reading Mastery*

Jean Stockard, Ph.D. & Kurt Engelmann, Ph.D.

Abstract

Data from two different sites were used to examine how exposure to a highly academic curriculum is related to growth in beginning literacy and early reading skills from kindergarten through the end of third grade. In one site students in one school used the Direct Instruction (DI) program, *Reading Mastery (RM)*, from kindergarten through grade 3, while students in a nearby school with similar demographic characteristics and entry skills had *Open Court*. In the other site, comparisons were made between one cohort that had a whole language kindergarten experience and began the *RM* program in first grade with two other cohorts who had *RM* throughout their K-3 career. In both sites, students exposed to *RM* had significantly greater growth in Nonsense Word Fluency scores from mid-kindergarten through the end of first grade. In addition, in both sites Oral Reading Fluency scores at the middle of first grade exhibited strong differences in favor of the *RM* students. For students in the Pacific Northwest site these differences persisted with very little change through the end of third grade. However, for those in the Midwestern site, where all cohorts had *RM* in grades 1-3, the differences gradually declined, although differences remained in favor of the *RM* group at the end of third grade.

Keywords: Direct Instruction, Reading Mastery, Educational Reform, and Academic Success

Introduction

A large body of literature has documented the relationship of early reading achievement to later academic accomplishments and economic and social well-being. Students who are poor readers in first grade have substantially higher probabilities of later academic, economic, and social problems than students who achieve at grade level at that time (e.g. Francis, Stuebing, Shaywitz, Shaywitz, & Fletcher 1996; Juel, 1988; Lipson & Wixson, 1997; Snider & Tarver, 1987; Wharton-McDonald, Pressley, & Hampston, 1998). These consistent and strong research findings have prompted extensive policy attention to promoting first grade reading achievement. One central concept in the literature is “readiness to learn,” the notion that all children should enter elementary school with the skills that prepare them to learn primary level academic content.

In the United States, kindergarten, literally translated as “children’s garden,” has traditionally been seen as the form of education that helps students transition from home to formal schooling and prepares them for the first grade academic experience. Kindergarten is now part of the public school system and universally available to all students in most jurisdictions in the United States. Yet a great deal of variability remains with regard to learning goals and curriculum. While definitions of what children should learn in the later grades of elementary schools are relatively standardized (even if the mode of teaching is not), definitions of what children should learn in kindergarten can vary from one jurisdiction to another and even from one school and one teacher within a school to another. A major component of this variation is the extent to which academic learning is emphasized within the curriculum, reflecting a division between those who emphasize minimal academic expectations with a “child-centered” approach and those who emphasize direct teaching of academic skills and content with a “teacher-directed” approach.

In the sections below we first examine literature regarding the importance of a teacher-directed approach to promote early academic preparation and progress in reading and then examine research on Direct Instruction (DI), a long-established teacher-directed approach and the focus of our study.

The Importance of Early Academic Preparation and Progress

The long-term impact of early academic learning has been captured with discussions of the “Matthew effect,” using the Biblical quotation that the “rich get richer and the poor get poorer” to describe the long-term and cumulative effects of good or poor reading skills on later academic success. A large body of empirical evidence demonstrates that early reading ability has lasting impacts on students’ academic careers. Those who are able to read fluently in first grade have much more success throughout their school careers. Early reading fluency results in exposure to much greater volume of material, and thus also produces a strikingly greater accumulation of vocabulary, language skills and bodies of knowledge (Cunningham & Stanovich, 1997, 1998; Francis, et al, 1998; Gough & Juel, 1991; Juel, 1988; Stanovich, 1986).

The importance of a teacher-directed approach in promoting success in early schooling has been supported by a relatively large body of literature and summarized by the National Reading Panel’s report on reading instruction. The report identified five areas of reading instruction that should be part of children’s primary grade instruction: phonemic awareness, phonics, fluency, vocabulary, and text comprehension. The panel, accompanying meta-analyses of the research literature, and numerous individual studies have demonstrated that phonemic awareness and phonics-oriented pre-literacy and early literacy instruction play a crucial role in enhancing early reading achievement (National Institute of Child Health and Human Development, 2000; Ehri, Nunes, Stahl, & Willows, 2001; see also Ball & Blachman, 1991; Blachman, Tangel, Ball, Black, & McGraw, 1999; Cavanaugh, Kim, Wanzek, & Vaughn, 2004; Foorman, Fletcher, Francis, Schatschneider, & Mehta, 1998; Kamps et al, 2007, 2008; Simmons, et al, 2007; Stuart 1999, Vandervelden and Siegel 1997).

The areas identified by the National Reading Panel’s report parallel theoretical models regarding the development of reading skills (e.g., Chall, 1983; Ehri, 2005; Ehri & McCormick, 1998; Simmons & Kame’enui 1998). These models describe how the foundational skill of phonological awareness, or being able to hear and manipulate sound structures, precedes the development of alphabetic understanding, or the understanding of the relation of print to speech. This, in turn, precedes phonological recoding of letter strings to sounds, which precedes the eventual reading of words and then connected text. The various models see these skills as overlapping, but ranging along a continuum, with the end goal of attaining fluency in reading by the end of the primary grades. Notably, this time point (grade 3) is also when the first high-stakes assessment is administered in U.S. schools and corresponds to the national goal, established within the No Child Left Behind Act, that all children will read by the end of grade 3 (Good, Simmons, & Kame’enui 2001).

As described by Fuchs and colleagues (Fuchs, Fuchs, Hosp, & Jenkins, 2001), oral reading fluency “is the oral translation of text with speed and accuracy” and “represents a complicated, multi-faceted performance” and a “complex orchestration,” (pp. 239-40) where students can read text in a fluid, automatic and seemingly effortless manner, thus allowing their intellectual efforts to be more directed toward comprehension than decoding. Extensive research has demonstrated that greater oral reading fluency is strongly related to better reading comprehension (Fuchs, et al, 2001; see also Baker, et al. 2008).

Because the process of developing reading fluency is developmental, occurring throughout the early school years, researchers have developed increasingly sophisticated methods of measuring this development through indicators of children’s growth in skills. This work builds on the curriculum-based measurement methodology (CBM) of Deno Deno, Mirkin, and Chiang (1982). The CBM methodology originally focused on oral reading fluency, but has now expanded to include measures of skills in the earlier stages of the developmental reading process, such as recognizing letter names and sounds. Two of the most commonly used systems are the *Dynamic Indicators of Basic Early Literacy Skills* (DIBELS,

2008; Good, Simmons, & Smith, 1998; Kaminski &Good, 1996) and *AIMSweb* (AIMSweb 2009; Fuchs & Fuchs, 1986). Both systems incorporate assessments of various elements of reading development including children's ability to recognize letters and link sounds and letters. The most important aspect of all of these systems is regular, systematic, efficient assessment of children's skills, with repeated short testing sessions during the school year, and comparison of these assessments to established benchmarks that indicate if children are making the progress needed to achieve the goal of reading fluently by the end of grade 3 (Good, Simmons, & Kame'enui, 2001). A number of school systems have adopted these measures as ways to monitor children's progress and better assess when additional help is needed.

To summarize, research has identified the skills that children need to acquire to learn to read fluently by the end of third grade. Research indicates that the development of strong reading ability builds on foundational early literacy skills. Assessment tools are available to monitor the development of these early skills as well as gains in reading fluency. Our research uses these assessment tools to examine the efficacy of one curricular program, *Reading Mastery (RM)*, in promoting growth in these early skills and the development of reading fluency.

Early Education and the Direct Instruction Model

Four decades before the report of the National Reading Panel was issued and two decades before the development of curriculum-based measurement, the DI model, which is the focus of our study, was developed. The original model was based on work with preschoolers in an “at-risk” population (Engelmann, 2007). Since that time the model has expanded to include a wide variety of curricular programs appropriate for multiple ages and grade levels and different subject areas. Yet, full implementation of the program calls for extensive academic work in kindergarten, or the preschool years, seeing this early period as key to a solid start in school and, especially, for catching at-risk children up to their more advantaged peers. It also involves continual assessment of children’s skills with adjustment of instruction to promote the highest achievement possible. Thus, the original design of the DI programs embodied the recommendations of the National Reading Panel regarding the crucial elements of reading instruction as well as the underlying notions of curriculum based assessment.

All of the DI programs seek efficiency and effectiveness of instruction through program design, organization of instruction, and positive student-teacher interaction. The DI approach attempts to control all the major variables that impact student learning through the placement and grouping of students into instructional groups, the rate and type of examples presented by the teacher, the wording that teachers use to teach specific concepts and skills, the frequency and type of review of material introduced, the assessment of students’ mastery of material covered and the responses by teachers to student’s attempts to learn the material.

DI programs are constructed according to a small step design that teaches isolated skills and concepts in separate tracks that are systematically integrated with skills and concepts in other tracks in increasingly sophisticated applications. For this reason, lessons do not focus on a single skill or topic. Instead, only about 10% of a lesson’s contents are new. The rest of the lesson is devoted to reviewing and applying skills and concepts that were introduced in previous lessons. Placement in the program is a critical factor in the program’s success. A major goal of DI is to build students’ confidence in their ability to learn while they master key skills and concepts. As Gersten, Darch, and Gleason (1988) note, “perhaps the central image that guided the conceptualization of Direct Instruction kindergarten was the image of students learning new concepts and skills each day, but in such a way that they experienced unremitting success” (p. 229). Placement at the point in the program in which students have already mastered material previously covered allows them to experience such success.

A large body of research has documented the efficacy of the general body of DI programs (e.g., Adams & Engelmann, 1996; Borman, Hewes, Overman, & Brown, 2003 ; Crowe, Connor, & Petscher, 2009), and several studies have specifically examined achievement growth of students who began receiving DI programs in kindergarten. For instance, Kamps et al. (2003) followed students from kindergarten through second grade, comparing growth in reading skills, using the DIBELS measures, of students receiving DI's *RM* with those in two other programs. While they found that students' skills grew over time with all three programs, growth was greatest for students exposed to *RM*.

Carlson and Francis (2002) also looked at changing achievement from kindergarten through second grade, comparing changes in scores on standardized achievement tests of students exposed to *Reading Mastery* and those exposed to comparison curriculum. They found that the students exposed to *RM* beginning in kindergarten had higher achievement scores than students in the comparison curriculum at both the end of first grade and the end of second grade. Through statistical analyses they determined that the pace of growth in first grade was significantly stronger for those in *RM* than in the comparison curriculum. However, during the second grade year the changes in achievement were similar for the two groups. Thus, the achievement advantage of the *RM* students reflected their greater achievement gains in kindergarten and first grade.

Comparable results have been reported for growth from kindergarten through the end of first grade for samples of at-risk students. Two studies (Gunn, Biglan, Smolkowski, & Ary, 2000; Kamps, et al, 2008;) used DIBELS measures to examine growth in reading achievement, and both found that students exposed to DI had significantly greater growth than those using other curricula. Similar results were found with analyses of scores on standardized achievement tests. Although the intervention in the study reported by Gunn et al only lasted for 2 years, follow-up analyses one year after the end of intervention indicated that the significant differences in achievement persisted both one and two years after instruction ceased (Gunn, Smolkowski, Biglan, & Black, 2002; Gunn, Smolkowski, Biglan, Black, & Blair, 2005).

We were able to find only one other study of the impact of DI that followed children from kindergarten through third grade. Gersten, Darch, and Gleason (1988) examined data from one community involved in Project Follow Through, a large educational experiment conducted in 20 communities throughout the country from 1969 to 1977 (Becker et al, 1981; Stebbins, Et. Pierre, Proper, Anderson, & Cerva, 1977). They compared the achievement of students in two different cohorts – one that began DI in kindergarten and another that began DI in first grade – to demographically similar students in the same district who had the district's traditional curriculum. The authors found that students who started school in first grade and received three years of instruction with DI significantly outperformed comparison students who received the district's program in mathematics and language, but not in reading comprehension or vocabulary, on the Metropolitan Achievement Test (MAT). In contrast, children who started DI in kindergarten significantly outperformed the comparison group when they reached third grade in reading comprehension as well as mathematics and language. The cohort that received DI in kindergarten scored near the national median in all measures in third grade, while those who received DI only in first through third grades had significantly lower scores. Paralleling the results of Gunn et al (2002, 2005), they found that the advantages accruing to the students in the DI kindergartens persisted after the program ended, with significant differences in reading achievement appearing through a final available data period at ninth grade.

To summarize, comparisons of growth in reading skills of students who have had kindergarten instruction with the DI model with those with other curricula indicate that students receiving DI have greater rates of growth and higher achievement. However, we found only one study that examined growth of students receiving instruction through the end of the primary grades (K-3), and this study involved analysis of data collected more than 30 years ago. Over the ensuing decades, while retaining many of its

original characteristics, the DI programs have been expanded and modified, and it is important to determine to what extent the earlier findings (both of the superiority of receiving DI in kindergarten and the extent to which beginning DI in first grade fails to compensate for this experience) can be replicated. In addition, the earlier studies provide somewhat conflicting evidence on the manner in which the advantages accrue to students in the DI programs. Are the advantages primarily due to acceleration of DI students in the earliest grades, as found by Carlson and Francis (2002)? Or do the different trajectories of growth continue over time, leading to widening gaps, as found by Gersten and associates (1988)?

The current study addresses these issues by examining gains in beginning literacy skills and oral reading fluency from kindergarten through third grade for students who received kindergarten instruction in *RM*, a highly structured DI academic program, and those who experienced less structured programs. Our analysis focuses on several interrelated questions: Do children exposed to *RM* have greater gains in beginning literacy skills and reading fluency than children in other programs? When do these differential gains occur and how long do the effects last? Can receiving a strong academic program in the primary grades (1-3) reduce the differential rates of achievement?

METHOD

Participants and Setting

Our first data set was from a K-12 district in the Pacific Northwest that is on the outskirts of a medium-sized city. The district has five elementary schools and 5700 students in K-12, two of which participated in the study. One adopted the DI program, *RM*, as the core reading curriculum for the primary grades including kindergarten as well as for students identified for special education. The other school used *Open Court*, a curriculum that has received high ratings from the Florida Center for Reading Research (FCRR, 2004). The school also occasionally used DI programs for students that teachers felt would benefit from the instruction. Teachers in the control school were provided support in implementing *Open Court* through a dedicated curriculum specialist as well as district resources, such as school psychologists and special education teachers. However, they did not have systematic technical support or guidance in their implementation of DI programs, and their use of these materials was best termed as incidental and occasional. We were not able to ascertain which students in the comparison school received extra help with DI, so this could not be included in our analysis. However, inclusion of students who might have received DI in the comparison group would diminish the probability of having significant results in favor of *RM*, thus providing a conservative test of its impact.

Data were available for 168 students who were enrolled in their respective schools from kindergarten through third grade. Only students who were in the sample for the entire time range were used in the analyses. Almost ninety percent of the students were non-Hispanic whites (89%), slightly less than a third (29%) were eligible for free or reduced lunch, and one-fifth 20% were classified as eligible for special education. There were no significant differences between the students in the two schools in these variables.

The second group of participants was drawn from a small, rural Midwestern district. It serves students through grade twelve and has four elementary schools. Unlike the Pacific Northwest district, this system adopted DI district-wide. In 2004 they began implementing *RM* in all of the elementary schools. Before that time, they used a whole language approach and a variety of reading and language programs. We analyzed data from three cohorts: one that began kindergarten in 2003 (n = 104), and thus had the traditional kindergarten curriculum, and cohorts that began kindergarten in 2004 (n = 100) and 2005 (n=114), who were taught using *RM* beginning in the kindergarten year. Data were available through the end of third grade for the first two cohorts and through the end of second grade for all three cohorts. As

with the Pacific Northwest site, only students who were in the sample for the entire time range were used in the analyses.

The Midwestern district had relatively more students who would be deemed “at risk” than the Pacific Northwest site, based on socio-demographic characteristics. Approximately one-third of the students (32.8%) were racial-ethnic minorities, primarily Hispanic, and half (50.6%) qualified for free or reduced lunch. However, less than one-tenth (7.8%) of the students were classified as eligible for special education, substantially fewer than in the Pacific Northwest district. There were no significant differences between the three cohorts in these characteristics.

Intervention Curriculum

Both sites used the DI program, *RM*, with technical support in the form of on-going training and in-class observations and coaching provided by the National Institute for Direct Instruction (NIFDI). The kindergarten *RM* program first concentrates on oral language skills to ensure that students are familiar with basic directions and to develop students’ background knowledge. Students then receive instruction in the reading curriculum, which initially focuses on pre-reading skills (phonemic awareness and phonics). Students start to read words within 30 lessons of reading instruction. The *RM* program introduces the sounds letters make rather than the letters themselves since decoding does not depend on knowing the names of letters (Engelmann, 2004). Students are taught to blend together the sounds when initially decoding words, and later to read whole words “the fast way.” The program concentrates on accuracy of decoding before fluency. Orthographic modifications used to introduce the sounds are gradually removed in the second level of the program.

Measures

Both districts administered the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) to all students in kindergarten and the primary grades at the times specified by the DIBELS guidelines (DIBELS, 2008). We used two measures within the DIBELS system obtained at the beginning of kindergarten as indicators of children’s initial skill levels: Letter Naming Fluency (LNF), which assesses children’s ability to correctly identify letters of the alphabet and Initial Sounds Fluency (ISF), which assesses children’s skills at identifying and producing the beginning sounds of a word. Two other measures were used as indicators of children’s reading development: Nonsense Word Fluency (NWF), which measures the ability to read phonetic nonsense words, which was assessed from the middle of kindergarten through the end of first grade (five testing periods); and Oral Reading Fluency (ORF), which measures the rate at which children can correctly read connected text in grade-level materials and was assessed from the middle of first grade through the end of third grade (eight testing periods).

Because the connected text used for the measure of oral reading fluency is taken from grade-level material, comparisons of ORF scores from one year to the next may not provide the most optimal picture of changes in skills over time. In other words, an ORF reading score for grade 1 is not directly comparable to one for grade 3 because the two tests use different reading material. To compensate, we transformed the ORF scores into Lexiles, a developmental scale of reading that ranges from less than zero for those who are just beginning to read to above 1700L for advanced readers. Thus, it adjusts for the different content used in the ORF at each grade level. The equations used for the conversions were developed from an extensive study involving over 2000 children in grades K-3 and several dozen reading passages (MetaMetrics 2009).

Analyses

To ensure that students within the groups were equivalent, we began our analyses by comparing their LNF and ISF scores at the first testing in the kindergarten year. We also compared the socio-demographic characteristics of the groups and, for the Pacific Northwest site, the achievement scores on statewide assessments for years before the study began. T-tests and effect sizes (Cohen's d) were used to compare the scores of the two groups.

We then examined changes in reading over time using linear growth modeling (Raudenbush & Bryk, 2002; Singer, 1998). We first looked at growth in nonsense word fluency from mid-K through the end of first grade using time (both the linear and quadratic effect), initial LNF and ISF scores, group, and the interaction of time and kindergarten experience as predictors. For the Pacific Northwest site, a dummy variable for group distinguished students in the *RM* school from students in the other school. For the Midwestern site, we used dummy variables to distinguish students in the three cohorts, allowing us to control for the possibility that teachers' greater experience with *RM* (for cohort 3) could be related to achievement differences.

In our analyses of growth in NWF seven models, each incrementally more complex than the previous one, were examined: 1) an intercept only model to provide a baseline indication of variation in NWF over the five time points, 2) a model that added the linear effect of time, 3) a model that added the quadratic effect of time, 4) a model that added LNF and ISF scores at the beginning of kindergarten, 5) a model that added dummy variable for group, 6) a model that added the interaction of group and the linear influence of time, and 7) a model that added the interaction of group and the quadratic influence of time.

To examine changes in ORF from mid-first grade through the primary years we also used linear growth modeling, employing models identical to those used in the examination of NWF. The Lexile scores were regressed on time (both linear and quadratic); LNF and ISF scores at the beginning of kindergarten; group, again using dummy variables; and the interactions of time and group.

As a way of illustrating the results, we graphed the growth in average NWF and ORF Lexile scores for each group. In addition, we computed effect sizes comparing the scores of students in the two groups within each site at both the beginning and end points of the series. For these calculations we used the standard formula for Cohen's d (difference of means divided by the common standard deviation). Finally, we used the descriptive data to estimate the cumulated differences in reading volume experienced by children within the different groups. This was calculated by simple estimates of words read over the course of a school year for students with average mid-year ORF scores at each grade level in each group.

To summarize, we examined our major question, do children exposed to the DI program, *RM*, have greater gains in beginning literacy skills and reading fluency than children in other programs, with linear growth models, graphs of the results, and calculations of effect sizes and volume of reading experience. Examination of these models and results allows us to see when these differential gains occur, if differences remain through the end of the primary grades, and whether beginning *RM* in the primary grades (1-3), as occurred in the Midwestern site, can compensate for differentials that may appear in the kindergarten years and early first grade. Thus, across the two sites we compare three different curricular experiences: 1) having *RM* as the core reading curriculum from kindergarten through third grade (a school in the Pacific Northwest site and cohorts 2 and 3 in the Midwest site), 2) having *Open Court* as the core reading curriculum from kindergarten through third grade (a school in the Pacific Northwest site), and 3) having a whole language program in kindergarten but *RM* in grades 1-3 (cohort 1 in the Midwest site).

RESULTS

Initial Differences between Groups

Tables 1 and 2 provide comparative information for students and schools in each group and at each site. DIBELS scores at the beginning of the kindergarten year are in the first panel of each table. Preliminary analysis indicated that there were minimal differences in the patterns of results for cohorts 2 and 3 in the Midwestern site. In other words, the statistical and substantive conclusions presented here are virtually identical to those that appear when data for cohorts 2 and 3 are distinguished. Thus, to simplify our presentation of the data we collapsed data for these two cohorts for our final analyses, using one dummy variable to distinguish those who received *RM* in kindergarten from those who did not. Results with the three cohorts separated (that is, with two dummy variables for group rather than one) are available upon request from the authors.

As would be expected given their different socio-demographic characteristics, students in the Pacific Northwest site (Table 1) had markedly higher LNF and ISF scores than those in the Midwestern site. Within each site there were only small differences in scores between the groups of students and none were statistically significant. Three of the differences favor the groups that did not have *RM*, while one (ISF in the Pacific Northwest site), favors the *RM* group. Thus, at baseline, there was little indication of differences in initial skills between the students who were in the more academically oriented kindergartens and exposed to *RM* in kindergarten and the other students. It should be remembered, however, that, to provide additional rigor to our results, we included these scores as control variables in the multivariate analyses.

Table 1. Comparison of Control School and DI School, Pacific Northwest Site

<i>A. Skills at Beginning of Kindergarten</i>					
	Control	DI	t	prob.	Cohen's d
	M (s.d.)	M (s.d.)			
Letter Naming Fluency	17.8 (13.0)	16.1 (13.5)	.84	.40	.13
Initial Sound Fluency	14.9 (10.0)	17.4 (12.5)	-1.47	.14	-.22
<i>B. Socio-Demographic Characteristics</i>					
	Control	DI	t	prob.	d
	%	%			
Non-Hispanic White	87	90	-.70	.48	-.11
Special Education	19	20	-.23	.82	-.04
Free and Reduced lunch	27	31	-.55	.58	-.09
<i>C. Percentage of 3rd to 5th Graders Meeting or Exceeding State Reading Benchmarks</i>					
	Control	DI	t	prob.	d
	%	%			
1998 - 1999	77	66	-2.62	0.01	-0.24

1999 - 2000	80	70	-2.49	0.01	-0.23
2000 - 2001	77	71	-1.48	0.14	-0.14
2001 - 2002	86	69	-4.36	<.0001	-0.40

Note: Instruction in DI began in the 2002-03 school year. All probability values are for two-tail tests. There were 85 students in the control school and 84 students in the DI school. Approximately 70 students were in each grade for the state-wide testing, with over 95 percent of all students participating in the testing in each year.

Table 2 Comparison of Cohorts, Midwestern Site

<i>A. Skills at Beginning of Kindergarten</i>					
	Cohort 1	Cohorts 2 & 3		prob.	Cohen's d
	M (s.d.)	M (s.d.)	t		
Letter Naming Fluency	11.9 (12.9)	11.0 (12.2)	.63	.53	0.07
Initial Sound Fluency	9.0 (8.2)	8.4 (7.5)	.73	.46	0.08

<i>B. Socio-Demographic Characteristics</i>					
	Cohort 1	Cohorts 2 & 3		prob.	d
	%	%	t		
Non-Hispanic White	72	64	1.47	.14	.17
Special Education	11	7	1.02	.31	.13
Free and Reduced Lunch	42	50	-1.29	.20	-.15

Note: All probability values are for two-tail tests. There were 104 students in Cohort 1, which did not have Reading Mastery in kindergarten, and 214 students in cohorts 2 and 3, which did have Reading Mastery in kindergarten.

Tables 1 and 2 also include details on the socio-demographic characteristics of the students. As noted above, that there were no significant differences in these characteristics between the groups at either site. Finally, Table 1 includes information from the Pacific Northwest site on the percentage of upper elementary students (grades three to five) who met or exceeded state reading benchmarks in the four years immediately preceding the beginning of the study. This was the only school wide information available for that period, but allows us to examine the possibility that there were pre-existing differences in school level achievement patterns. Differences were statistically significant in three of the four years, but favor the control school. In other words, the historical data pattern indicates that by the upper primary grades the students in the school that implemented *RM* had significantly fewer students meeting or exceeding the state reading benchmarks. Such a difference would be considered conservative in nature, making it even less likely that results in our analysis would favor the *RM* group.

Growth in Nonsense Word Fluency

Table 3 summarizes the analyses of growth in nonsense word fluency from mid kindergarten through the end of first grade. The -2 Log-Likelihood Ratio statistics, which have a chi-square distribution, were used to determine the models that best fit the data. (Details on model fit statistics are available from the authors on request.) For the Midwestern site, the most complex model (Model 7), which included time, the quadratic term for time, initial skills, group, and the interactions of group and both the linear and quadratic effect of time, provided the best fit. For the Pacific Northwest site, however, this model was not significantly better than Model 6, which omitted the interaction of group and the quadratic term for time. In other words, for both data sets, the best fitting models were those that indicated that the changes over time in NWF varied between students who had *RM* and those who did not, but the precise nature of these differences in growth varied between the two sites.

Table 3 Growth Curve Analysis of Nonsense Word Fluency from Mid-Kindergarten Through the End of First Grade by Site

	Fixed Effects					
	Pacific Northwest Site			Midwestern Site		
	b	s.e.	prob.	b	s.e.	prob.
Intercept	13.8	2.3	<.0001	18.3	1.9	<.0001
Time	-0.2	1.6	0.90	9.0	2.0	<.0001
Time squared	3.5	0.3	<.0001	0.1	0.4	0.89
LNF - Start of K	0.8	0.1	<.0001	0.6	0.1	<.0001
ISF - Start of K	0.2	0.1	0.04	0.2	0.1	0.07
Group (RM in Kindergarten)	----	----	----	7.1	2.0	0.0005
Group (RM School)	-3.8	2.1	0.08	----	----	----
Time * Group	2.9	1.3	0.03	-3.4	2.4	0.16
Time Squared * Group	----	----	----	1.8	0.5	0.0008
Random Effects						
	var.	s.e.	prob.	var.	s.e.	prob.
Var between persons (11)	39.4	22.5	0.04	25.3	18.1	0.08
Var within persons (12)	19.1	9.9	0.053	37.0	7.0	<.0001
Var in growth rates (22)	46.3	7.9	<.0001	31.2	5.0	<.0001
Residual	242.2	15.4	<.0001	294.3	13.5	<.0001

Note: t values for coefficients may be calculated by dividing the coefficients by the standard errors (s.e.). Probabilities for the fixed effects adjusted for different degrees of freedom for the time-varying variables and the time-invariant variables.

Results for the Pacific Northwest site are given in the left-hand columns of Table 3, and those for the Midwestern site are in the right-hand columns. Random effects coefficients for each model are in the bottom rows of Table 3. These coefficients are significantly greater than zero for the residual term, for the variance in growth rates, and for the within-person variance. For the variance between persons, the term is

marginally insignificant for the Midwestern site ($p = .08$), but significant for the Pacific Northwest site ($p=.04$). These results indicate that the variables in our model were not sufficient to account for all of the variance in students' scores and changes in these scores over time.

The top rows of Table 3 give the fixed effect coefficients for the best fitting models for each site. To facilitate interpretation of the coefficients, time was coded with 0 as the score for the first observation (mid-kindergarten). Thus, the intercept term indicates the expected value of NWF at that time, and the values for time and time squared indicate the expected changes in score from one testing period to the next. The positive coefficients associated with the LNF and ISF scores (measured at the start of kindergarten) indicate that students with higher initial skills had higher expected NWF scores. As hypothesized, the interaction of group and time was statistically significant for both sites, with students who had *RM* throughout their academic career having significantly higher rates of growth in NWF over time than other students. However, as noted above, the exact pattern of these differences varied between the two sites.

With the Pacific Northwest sample, all students had an accelerated pattern of growth in NWF over time, as indicated by the positive coefficient associated with the quadratic term for time. However, the linear effect of time differed significantly between the two groups, with a value near zero (-0.2) for students who did not have *RM*, to a value close to three words per testing period ($+2.7 = -0.2 + 2.9$) for those exposed to *RM*. The result was a widening gap in the NWF scores of students in the two groups. This is shown in Figure 1, which graphs the mean values of NWF at each time point for students in each group. Students in the *RM* school had lower scores than those in the other school at the initial testing, but this pattern reversed by the end of first grade. The differences are also illustrated by simple calculations of Cohen's *d* (the difference of means divided by the common standard deviation). At the first testing period, in the middle of kindergarten, $d = -.21$ ($= (25.7-29.5)/18.3$), reflecting the higher scores of students in the control school. By the last testing period at the end of first grade, $d = .24$ ($= (95.1-85.6)/39.8$), indicating an educationally important advantage for the *RM* students and illustrating their more rapid gains.

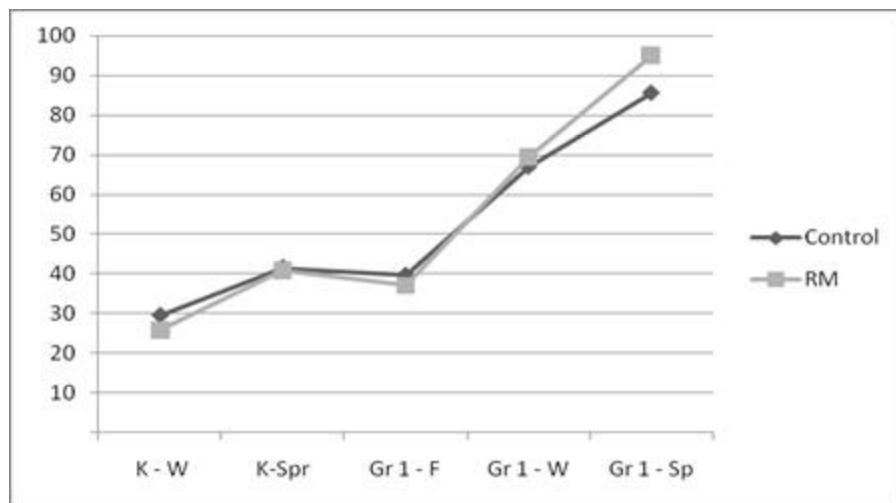


Figure 1. Growth in Nonsense Word Fluency, Mid-Kindergarten to Spring of First Grade, by Group, Pacific Northwest Site.

With the Midwestern sample, the main effect of having *RM* in kindergarten was significant, indicating that these students already had higher expected NWF scores by the first testing point (even

though they had started school with equivalent scores on LNF and ISF). The interaction of *RM* in kindergarten and the linear effect of time was negative, but insignificant, while the interaction of having *RM* in kindergarten and the quadratic effect of time was positive and highly significant ($p = .0008$). The net result was the same as in the Pacific Northwest site, with stronger growth over time in NWF for students who had *RM* throughout their academic careers. However, the quadratic effect indicates that this was more apparent at the later time points. Figure 2 illustrates the differences between the cohorts and the widening gap between the two groups. For this site the values of Cohen's d always favored the students in the *RM* kindergarten, but, as in the Pacific Northwest site, the values became larger over time – moving from .42 ($=(29.7-23.7)/14.4$) at the middle of kindergarten to .56 ($=(86.2-64.9)/36.0$) at the end of first grade.

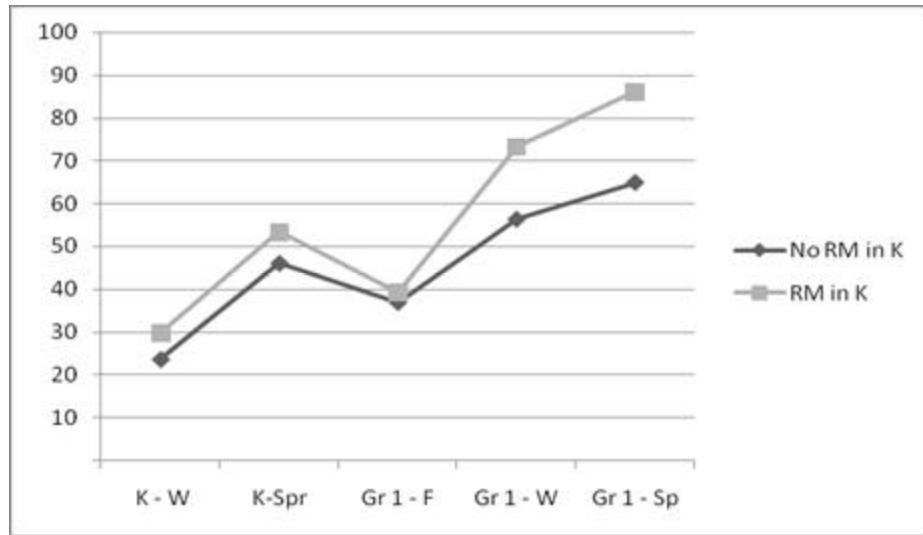


Figure 2. Growth in Nonsense Word Fluency, Mid-Kindergarten to Spring of First Grade, by Group, Midwestern Site

Growth in Oral Reading Fluency

Table 4 gives results of the analysis of growth in Oral Reading Fluency Lexiles from the middle of first grade through the end of third grade. The model fit statistics (available on request from the authors) indicated that the best fitting model for both sites was the most complex model (Model 7), which included time, the quadratic term for time, initial skills, group, and the interactions of group with both the linear and quadratic effects of time. In both sites, students who had higher skills at the beginning of kindergarten had significantly higher oral reading fluency scores in the primary grades (the significant effects associated with LNF and ISF). In addition, in both sites scores increased significantly over time (the positive effect associated with time), although there was also some deceleration in this growth over time (the negative effect associated with time squared). Most relevant to our hypotheses, students who had *RM* throughout their elementary schooling had significantly higher ORF scores than those who did not (the significant positive effect associated with group). The magnitude of this advantage was roughly similar between the two sites: 142 Lexiles in the Pacific Northwest site and 157 Lexiles in the Midwestern site – a difference at the first grade level of about 20 words per minute.

Table 4 Growth Models of Change in ORF (Lexiles) From Mid First Grade Through End of Third Grade by Site

	Fixed Effects					
	Pacific Northwest Site			Midwestern Site		
	b	s.e.	prob.	b	s.e.	prob.
Intercept	-345.4	33.6	<.0001	-285.9	25.7	<.0001
Time	202.2	6.9	<.0001	180.1	6.4	<.0001
Time squared	-12.6	0.9	<.0001	-10.3	0.9	<.0001
LNF - Start of K	10.1	1.3	<.0001	10.3	1.1	<.0001
ISF - Start of K	5.0	1.5	0.001	5.0	1.7	0.005
Group (RM in Kindergarten)	-----	-----	-----	156.9	26.2	<.0001
Group (RM School)	142.5	34.1	<.0001	-----	-----	-----
Time * Group	-17.8	9.8	0.07	-13.2	8.1	0.10
Time Squared * Group	2.2	1.3	0.08	-1.5	1.1	0.18
Random effects						
	var.	s.e.	prob.	var.	s.e.	prob.
Var between persons (11)	40109	5017	<.0001	38812	3577	<.0001
Var within persons (12)	-1261	547	0.02	-173	320	0.59
Var in growth rates (22)	594	96	<.0001	185	49	<.0001
Residual	10673	483	<.0001	12826	448	<.0001

Note: t values for coefficients may be calculated by dividing the coefficients by the standard errors (s.e.). Probabilities for the fixed effects adjusted for different degrees of freedom for the time-varying variables and the time-invariant variables.

In addition, however, the patterns of change over time differed between the groups. (While none of the interaction effects were statistically significant at standard levels, the addition of these terms to the models provided significantly better overall fit to the data, and it is, thus, important to discuss them.) For the Pacific Northwest site, where students in one school had *RM* throughout their primary years and the other students did not, the interaction of group with the linear effect of time was negative, while the interaction of group with the quadratic effect of time was positive. The net result was that, over time, the two interactions effectively canceled each other out. Figure 3 gives the average ORF Lexile scores at each testing for students in the two schools and illustrates how the advantage of students who received *RM* remained relatively constant throughout the primary years. Calculations of Cohen's d confirm this conclusion. Comparison of average scores of students in the two schools for the first testing at the middle of first grade resulted in an effect size of .46 ($= (16.6 - (-100.2)) / 275.8$), and the same comparison at the end of third grade resulted in an effect size of .42 ($= (877.6 - 767.4) / 264.2$). Both of these would be considered educationally meaningful.

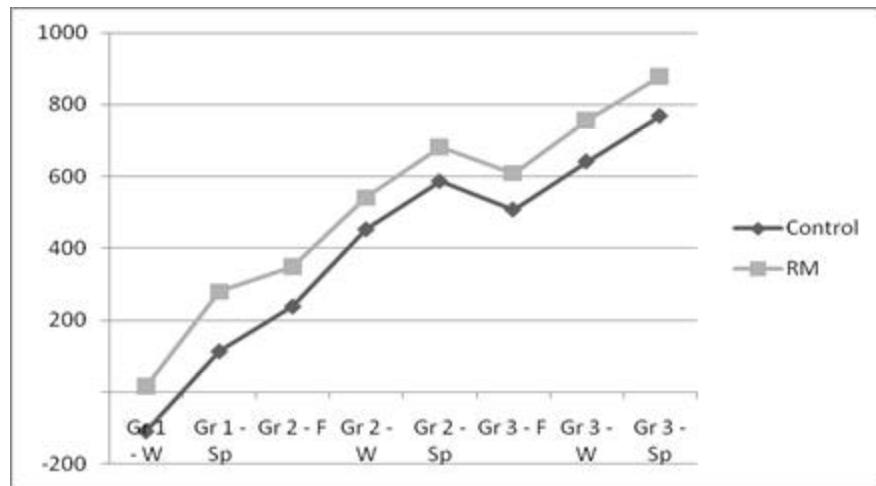


Figure 3. Growth in ORF Lexile Scores, Mid-First Grade Through Spring of Third Grade, by Group, Northwest Site

For the Midwestern site, where all cohorts had *RM* in the primary grades and the only difference was in the kindergarten curriculum, the results were slightly different. For this site, both interaction effects were negative. In other words, the growth in ORF after the beginning of first grade was slightly less for those who began *RM* in kindergarten than for those who did not have *RM* at that time. The students who were exposed to *RM* beginning in first grade were gradually catching up to the students who began *RM* in kindergarten. However, even with this faster growth, the students were not able to fully catch up, as illustrated in Figure 4. The students without the kindergarten *RM* experience always had lower average ORF Lexile scores than students who began *RM* while in kindergarten, although the differences diminished substantially over time. This pattern is illustrated by the calculations of Cohen's *d*, which declined from .53 ($= (36.1 - (-113.6)) / 282.83$) at the middle of first grade, indicating strong and educationally meaningful differences, to .10 ($= (722.1 - 695.6) / 268.8$) by the end of third grade, a level no longer considered educationally important.

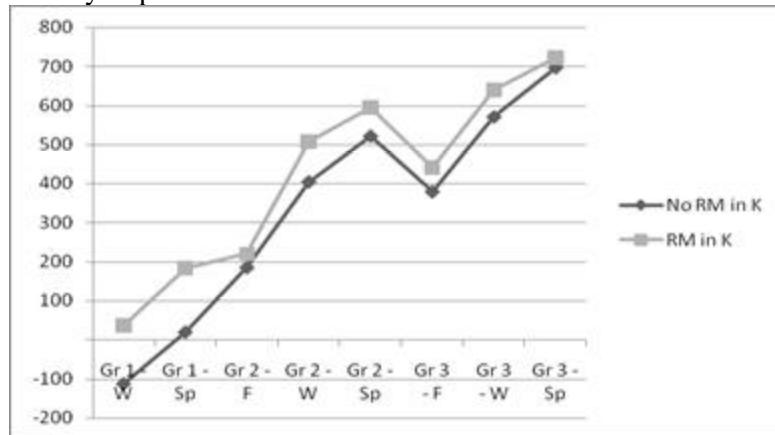


Figure 4. Growth in ORF Lexile Scores, Mid-First Grade Through Spring of Third Grade, by Group, Midwestern Site

Finally, Table 5 translates the differences in oral reading fluency into a measure of "reading volume," the number of words that an average child in each setting and group would be expected to read,

given their mid-year ORF score in each grade, if they read for 60 minutes a day throughout the school year. The calculations provide a simple way to estimate the cumulative impact of reading fluency and the associated exposure to enhanced language, vocabulary, and knowledge noted by reading scholars and discussed above. It can be seen that the differences are not trivial. In the Pacific Northwest site, where students in one school had *RM* from kindergarten through third grade and the other students did not, the average student in the *RM* school read from 720 to 960 more words a day, depending upon the grade level. In the Midwestern site, where all cohorts had *RM* in the primary grades and the only difference was in the kindergarten curriculum, the differences declined somewhat over time, but are still marked, ranging from over 1,000 words a day in first grade to 600 in third grade. The cumulative impact of these differences is shown in the last column. Assuming that the school year is 180 days, the differences translate into a discrepancy between the average students in the two groups, in each site, of almost one-half million words read throughout their primary grades.

Table 5. Estimated Differences between Groups in Words Read, by Grade and Site

		Estimated Differences in Words Read		
		Mid Year ORF		
		Control	DI	Daily
First grade		34	49	900
Second grade		87	99	720
Third grade		111	127	960
Total				464,400

		Estimated Differences in Words Read		
		Mid Year ORF		
		No DI in K	DI in K	Daily
First grade		34	51	1,020
Second grade		80	94	840
Third grade		101	111	600
Total				442,800

Note: Calculations for daily differences were derived by multiplying the difference of the two mid-year ORF scores in each grade by 60, assuming one hour of reading a day. The yearly difference was calculated by multiplying the daily score by 180, using this number as a typical length for a school year. Thus, calculations are estimates for average students in each site, assuming one hour of reading per day and limited to the 180 day school year.

DISCUSSION

This paper used data from two different sites to examine the relationship of academic kindergarten instruction, as well as variations in later instruction, to growth in beginning literacy and early reading skills of students from kindergarten through the end of third grade. In the Pacific Northwest site, students in one school had the highly structured DI program, *RM*, from kindergarten through the end of third grade. Their achievement growth was compared with that of students in a comparable nearby school in the same district that used *Open Court*, another highly rated, but somewhat less structured, curriculum throughout this time span. In the Midwestern site, comparisons were made between three

cohorts of students. Students in one cohort had a whole language kindergarten experience, but began using *RM* in first grade. Their achievement growth was contrasted with that of students in two other cohorts who had *RM* throughout their K-3 career. Both sites were dedicated to promoting higher achievement of their students, as indicated by their use of the DIBELS monitoring system and, in the Pacific Northwest site, by the use of a strong, phonics-based curriculum in the control school. As such, our analysis might provide a relatively stringent test of the efficacy of *RM* in promoting higher achievement.

Based on the literature reviewed above we expected that all students would have increased NWF scores from mid-kindergarten to the end of first grade, although it was unclear if a linear or a quadratic function would best fit the data. We also expected, given previous literature, that students with higher LNF and ISF scores at the beginning of kindergarten would have higher NWF scores throughout time (larger intercepts). Most important, we expected that the gains over time would be stronger for students with instruction in *RM*. In other words, we expected that there would be a significant time by group interaction and, potentially, a significant interaction between group and the quadratic term for time. With respect to the results with ORF scores, we expected that the intercepts would be higher for students who had *RM* throughout their school career (significant effects of group), reflecting stronger achievement gains for the *RM* students that occurred before the middle of the first grade year. As noted above, the literature provided less guidance as to differences between the slopes of the lines (the time by group interactions), although we expected that the advantage of having *RM* from the beginning of their schooling would be apparent through the end of third grade.

There were no statistically significant differences in measures of early phonological awareness (as indicated by letter naming and initial sound fluency measures) of students in the various groups at the beginning of kindergarten, and these measures were controlled in the multivariate analyses. In addition, there were no differences in the socio-demographic characteristics of students in the two groups at each site. Finally, the historical pattern of achievement in the Pacific Northwest site, which compared two different schools, indicated that, over time, upper grade students at the school that implemented *RM* were less likely than those in the control school to meet or exceed benchmarks. Thus, there was no indication that the *RM* students were advantaged by their initial skills, their socio-demographic characteristics, or the historical achievement pattern of their schools.

There were, however, significant differences between the groups in reading skills and patterns of growth over time. In both sites those who were exposed to *RM* throughout their academic career had significantly greater growth in Nonsense Word Fluency scores, resulting in significantly higher scores on this measure by the middle and end of first grade. Reflecting these differential growth patterns, there were also substantial, and statistically significant, differences between the groups, at each site, in initial Oral Reading Fluency scores at the middle of first grade. For students in the Pacific Northwest site these differences persisted with very little change through the end of third grade. However, for those in the Midwestern site, where all cohorts had *RM* in grades 1-3, the differences gradually declined, although they remained at the end of grade 3.

These results indicate that the advantage in reading fluency for the *RM* students was established by the middle of first grade. One could suggest that these differences reflect the variations in kindergarten experiences of the various groups. For the Midwestern site, where one cohort had *RM* in kindergarten and the other had a whole language approach, this explanation appears relatively easy to justify. Differences in skills, as measured by NWF, appeared by the middle of kindergarten and widened through first grade. For the students in the Pacific Northwest site, where the control students received instruction in *Open Court* throughout the K-3 years, the situation is somewhat different. The differences in NWF did not become strong until the first grade year. The *Open Court* curriculum has received strong ratings from various review bodies, and the lack of differences may reflect equal growth in skills in early stages of the

curriculum. The differences were stronger in the Midwestern site where the students in the first cohort had a curriculum that was less academic than *Open Court*.

However, this pattern may also be related to the way in which the measure of nonsense word fluency (NWF) relates to the *RM* curriculum. The NWF measure, which is the only measure of “reading” used in kindergarten, employs short vowel words, while the *RM* curriculum introduces short vowels only after long vowels are studied. Thus, the lack of strong differences on NWF scores between the groups in kindergarten for the Pacific Northwest site may reflect the extent to which the measure is accurately measuring what the children have learned. By first grade, when the *RM* children have been exposed to both short and long vowels, the differences become more apparent. In contrast, the ORF measure uses real words and is more likely to tap the skills that the children have learned.

As shown by the calculations in Table 5, the advantages that accrued to students with consistent instruction in *RM* were not trivial. Scholars using the concept of “reading volume” note the strong cumulative impact of increased exposure to printed material on general knowledge and language development. Children who read more are exposed to much richer vocabularies, more varied conceptual ideas, and a vastly broader range of information (Cunningham & Stanovich, 1990, 1998; Stanovich & West, 1989). Our computations of reading volume focused on the reading rate of the average student in each group and used minimal estimates of time engaged in reading (one hour a day and only in the school year). Fluent readers are more likely than other students to also read independently, and thus our estimates of the differential impact are probably conservative. The important point is that the differences in cumulative exposure to print were not small. In addition, these differences appeared in both sites. Even in the Midwestern schools, where students began work with *RM* in first grade and gradually began to catch up with the other cohort, there were strong differences in words read and the cumulative impact of the differences was only slightly smaller than in the other location.

Our results would also appear to provide support for the theories regarding reading development, which describe learning to read as a developmental process where reading fluency builds on the development of early phonological understanding and skills. Children’s growth in reading skills occurred throughout the time period of the study and was greater for those exposed to a more systematic and explicit curriculum whose logical ordering matches the theoretical formulation. Our results also replicate earlier studies that documented the superior achievement and growth in reading skills of students exposed to *RM*. Similar to findings of Carlson and Francis (2002), the acceleration of *RM* students in our samples was most apparent in the early grades. The differences in growth through kindergarten and first grade resulted in substantial differences in Oral Reading Fluency in mid first grade which persisted through the primary years (for the Pacific Northwest site) or only gradually declined (for the Midwestern site where all students were exposed to *RM* after grade 1). The advantage accruing to *RM* students appeared in comparison to a whole language kindergarten, but also in comparison to the *Open Court* curriculum, a finding that replicates other results (e.g. Crowe, Connor & Petscher, 2009; O’Brien and Ware, 2002, p. 191).

In short, as with the earlier studies, our more contemporary data indicate that students who had *RM* throughout their K-3 years were significantly more likely to be on track for academic success in the primary grades. By the middle of first grade, students who began studying with *RM* in kindergarten had significantly greater oral reading skills. These differences persisted through the end of the primary grades. Scholars have long emphasized the cumulative impact of early reading fluency and our simulations indicate that the differences in reading volume between our study groups were not trivial.

We believe that the results presented in this paper have implications for practitioners and policy makers. They provide additional support for the advocates of early academic instruction. In addition, they replicate earlier research indicating the effectiveness of *RM* in these academic kindergartens. Children who were provided with this instruction were significantly more likely than other students, even those

who began *RM* at first grade or who studied in another highly rated curriculum, to be on a trajectory toward continued academic success in later years.

That said, continued research in this field is, of course, important and can address limitations in our work. Our analysis was limited to the DIBELS measures and other work might include a wider variety of assessments, including norm and criterion referenced instruments. In addition, both implementations of DI in this study involved technical support from the National Institute for Direct Instruction, and it could be informative to understand more about how this technical support enhanced the success of the implementations. While the comparison groups had implementation support from district and school personnel that was equal in duration and expense, we had no way of gauging the quality of such assistance.

It is also important to include multiple sites in studies of curricular programs. Our examination of two different sites allowed for the replication of findings and results. For instance, the differences found in ORF scores at the middle of first grade between the groups were very similar at the two sites. In addition, however, the slight difference in design between the two sites allowed us to examine a more nuanced question of the extent to which beginning *RM* in first grade could alter the trajectory of achievement gain over time. Other work could include even more sites and examine the extent to which our results could be replicated with a variety of different populations.

Finally, our study did not, of course, involve random assignment of students to groups, a feat that would have been virtually impossible in the setting and time span employed in our study. Implementing some type of random assignment in a replication could provide additional controls. On the other hand, while experimental studies that randomly assign students to short-term curricular interventions might provide an aura of scientific control and internal validity, they can differ so radically from the real-life situations of schools that they have relatively low external validity. The long-term growth of achievement and the maintenance of high level skills is undoubtedly the most important goal of education. We suggest that this growth can only be thoroughly understood through long-term studies in real life settings.

REFERENCES

Adams, G. L., & Engelmann, S. (1996). *Research on Direct Instruction: 25 years beyond DISTAR*. Seattle, WA: Educational Achievement Systems.

AIMSweb (2009). AIMSweb: Progress monitoring and RTI system. <http://www.aimsweb.com/>. Accessed August 4, 2009.

Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'enui, E. J., et al (2008). Reading fluency as a predictor of reading proficiency in low-performing, high poverty schools. *School Psychology Review*, 37 (1), 18-37.

Ball, E. W. & Blachman, B. A. (1991). Does phoneme awareness training in kindergarten make a difference in early word recognition and developmental spelling? *Reading Research Quarterly*, 26, 49-66.

Becker, W. C., Engelmann, S., Carnine, D., & Rhine, R. (1981). Direct Instruction models. In R. Rhine (Ed.), *Encouraging Change in American Schools: A Decade of Experimentation* (pp. 95-154). New York: Academic Press.

Blachman, B. A., Tangel, D. M., Ball, E. W., Black, R. & McGraw, C. K. (1999). Developing phonological awareness and word recognition skills: A two-year intervention with low-income, inner-city children. *Reading and Writing: An Interdisciplinary Journal*, 11, 239-273.

Borman, G. D., Hewes, G. M., Overman, L. T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research* 73 (2), 125-230.

Carlson, C. D., & Francis, D. J. (2002). Increasing the reading achievement of at-risk children through Direct Instruction: Evaluation of the Rodeo Institute for Teacher Excellence (RITE). *Journal of Education for Students Placed At Risk*, 7(2), 141-166.

Cavanaugh, C. L., Kim, A., Wanzek, J., & Vaughn, S. (2004). Kindergarten reading interventions for at-risk students: Twenty years of research. *Learning Disabilities: A Contemporary Journal*, 2 (1), 9-21.

Chall, J. S. (1983). *Stages of reading development*. New York: McGraw-Hill.

Crowe, E. C., Connor, C. M., & Petscher, Y. (2009). Examining the core: Relations among reading curricula, poverty, and first through third grade reading achievement. *Journal of School Psychology*, 47, 187-214.

Cunningham, A.E. & Stanovich, K.E. (1990). Assessing print exposure and orthographic processing skill in children: A quick measure of reading experience. *Journal of Educational Psychology*, 82, 733-740.

Cunningham, A. E. & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33 (6), 934-945.

Cunningham, A. E. & Stanovich, K. E. (1998). What reading does for the mind. *American Educator*, (Spring/Summer), 1-8.

Deno, S. L., Mirkin, P., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children*, 49, 36-45.

DIBELS (2008). DIBELS data system: Using data to improve achievement for each and all. <https://dibels.uoregon.edu/>, accessed December 17, 2008.

Ehri, L. C. (2005). Learning to read words: Theory, findings, and issues. *Scientific Studies of Reading*, 9 (2), 167-188.

Ehri, L. C. & McCormick, S. (1998). Phases of word learning: Implications for instruction with delayed and disabled readers. *Reading and Writing Quarterly*, 14, 135-163.

Ehri, L. C., Nunes, S. R., Stahl, S. A., & Willows, D. M. (2001). Systematic phonics instruction helps students learn to read: Evidence from the National Reading Panel's meta-analysis. *Review of Educational Research*, 71 (3), 393-447.

Engelmann, S. (1969). *Preventing failure in the primary grades*. New York: Simon and Schuster.

Engelmann, S. (2000). About reading—A comparison between *Reading Mastery* and *Horizons. Effective School Practices*, 18(3), 15–26.

Engelmann, S. (2004). Foreword. In N. E. Marchand-Martella, T. A. Slocum, R. C. Martella (Eds.), *Introduction to Direct Instruction* (pp. xix-xxvi). Boston: Pearson Education, Inc.

Engelmann, S. (2007). *Teaching needy kids in our backward system: 42 Years of trying*. Eugene, Oregon: ADI Press.

Florida Center for Reading Research (FCRR). (2004). *Open Court Reading Pre-K*. Florida Center for Reading Research. <http://www.fcrr.org/index.htm>, accessed January 29, 2009.

Foorman, B. R., Fletcher, J. M., Francis, D. J., Schatschneider, C., & Mehta, P. (1998). The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology*, 90 (1), 37-55.

Francis, D. J., Stuebing, K. K., Shaywitz, S. E., Shaywitz, B.A., & Fletcher, J. M. (1996). Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology*, 88 (1), 3-17.

Fuchs, L. S., & Fuchs, D. (1986). Curriculum-Based Assessment of progress toward long-term and short-term goals. *Journal of Special Education*, 20, 69-82.

Fuchs, L. S., Fuchs, D., Hosp, M. K., & Jenkins, J. R. (2001). Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis. *Scientific Studies of Reading*, 5(3), 239-256.

Gersten, R., Darch, C., & Gleason, M. (1988). Effectiveness of a Direct Instruction academic kindergarten for low-income students. *Elementary School Journal*, 89(2), 227-40.

Good, R. H., Simmons, D. C., & Kame'enui, E. J. (2001). The importance and decision-making utility of a continuum of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading*, 5 (3), 257-288.

Good, R. H., Simmons, D. C., & Smith, S. B. (1998). Effective academic interventions in the United States: Evaluating and enhancing the acquisition of early reading skills. *Educational and Child Psychology*, 15 (1), 56-70.

Gough, P. B. & Juel, C. (1991). The first stages of word recognition. In: L. Rieben & C.A. Perfetti (eds.) *Learning to read: Basic research and its implications* (pp. 47-56). Hisllsdale, NJ: Erlbaum.

Gunn,B., Biglan, A., Smolkowski, K., & Ary, D. (2000). The efficacy of supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school. *The Journal of Special Education*, 34 (2), 90-103.

Gunn, B., Smolkowski, K., Biglan, A., & Black, C. (2002). Supplemental instruction in decoding skills for Hispanic and non-Hispanic students in early elementary school: A follow-up. *The Journal of Special Education*, 36 (2), 69-79.

Gunn, B., Smolkowski, K., Biglan, A., Black, C., & Blair, J. (2005). Fostering the development of reading skill through supplemental instruction. *The Journal of Special Education*, 39 (2), 66-85.

Hasbrouck, J. & Tindal, G. A. (2006). Oral reading fluency norms: A valuable assessment tool for reading teachers. *The Reading Teacher*, 59 (7), 636-644.

Kaminski, R. A. & Good, R. H. (1996). Toward a technology for assessing basic early literacy skills. *School Psychology Review*, 25, 215-227.

Kamps, D. M., Willis, H. P., Greenwood, C. R., Thorne, S., Laxo, J. F., Crocket, J. L., et al. (2003). Curriculum influences of growth in early reading fluency for students with academic and behavioral risks: A descriptive study. *Journal of Emotional and Behavioral Disorders*, 11 (4), 211-224.

Kamps, D., Abbott, M., Greenwood, C., Arreaga-Mayer, C., Wills, H. P., Longstaff, J., et al (2007). Use of evidence-based, small-group reading instruction for English language learners in elementary grades: Secondary-Tier intervention. *Learning Disability Quarterly*, 30, 153-168.

Kamps, D., Abbot, M., Greenwood, C., Wills, H. P., Veerkamp, M. & Kaufman, J. (2008). Effects of small-group reading instruction and curriculum differences for students most at risk at kindergarten: Two year results for secondary- and tertiary-level interventions. *Journal of Learning Disabilities*, 41 (2), 101-114.

Juel, C. (1988). Learning to read and write: A longitudinal study of 54 children from first through fourth grades. *Journal of Educational Psychology*, 85, 112-126.

Lipson, M. Y., & Wixson, K. K. (1997). *Assessment and instruction of reading and writing disability: An interactive approach* (2nd ed.). New York: Plenum.

MetaMetrics (2009). *Linking DIBELS oral reading fluency with the Lexile framework for reading*. Durham, NC: MetaMetrics, Inc.

National Center for Education Research. (2008). *Effects of Preschool Curriculum Programs on School Readiness: Report from the Preschool Curriculum Evaluation Research Initiative*. Retrieved on December 12, 2008 from <http://ies.ed.gov/ncer/pubs/20082009/index.asp>.

National Institute of Child Health and Human Development. (2000). Report of the National Reading Panel. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.

O'Brien, D.M. & Ware, A. M. (2002). Implementing research-based reading programs in the Fort Worth Independent school district. *Journal of Education for Students Placed at Risk*, 7, 167-195.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, California: Sage.

Schieffer, C., Marchand-Martella, N. E., Martella, R. C., Simonsen, F. L., Waldron-Soler, K. M. (2002). An analysis of the Reading Mastery program: Effective components and research review. *Journal of Direct Instruction*, 2(2), 87-119.

Simmons, D. C. & Kame'enui, E. J. (Eds.) (1998). *What reading research tells us about children with diverse learning needs: Bases and basics*. Mahwah, NJ.: Lawrence Erlbaum Associates, Inc.

Simmons, D. C., Kame'enui, E. J., Harn, B., Coyne, M. D., Stoolmiller, M., Edwards-Santoro, L. et al (2007). Attributes of effective and efficient kindergarten reading intervention: An examination of instructional time and design specificity. *Journal of Learning Disabilities*, 40 (4), 331-347.

Singer, J.D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics*, 24 (4), 323-355.

Snider, V. E., & Tarver, S. G. (1987). The effect of early reading failure on acquisition of knowledge among students with learning disabilities. *Journal of Learning Disabilities*, 20, 351–356.

Stanovich, K.E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21 (4), 360-407.

Stanovich, K.E. & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 24, 402-433.

Stebbins, L., St. Pierre, R. G., Proper, E. C., Anderson, R. B., & Cerva, T. R. (1977). *Education as experimentation, Vol. IV*. Cambridge, MA: Abt Associates.

Stein, M. & Kinder, D. (2004). Reading. In N. E. Marchand-Martella, T. A. Slocum, R. C. Martella (Eds.), *Introduction to Direct Instruction* (pp. 100-139). Boston: Pearson Education, Inc.

Stipek, D. J. & Ryan, R. H. (1997). Economically disadvantaged preschoolers: Ready to learn but further to go. *Developmental Psychology*, 33, 711-723.

Stuart, M. (1999). Getting ready for reading: Early phoneme awareness and phonics teaching improves reading and spelling in inner-city second language learners. *British Journal of Educational Psychology*, 69, 587-605.

U.S. Department of Education. (2008). *Reading First*. Retrieved on December 12, 2008 from <http://www.ed.gov/programs/readingfirst/>.

Vanderfelden, M. C. & Siegel, L. S. (1997). Teaching phonological processing skills in early literacy: A developmental approach. *Learning Disability Quarterly*, 20 (2), 63-81.

Wharton-McDonald, R., Pressley, M., & Hampston, J. (1998). Literacy instruction in nine first-grade classrooms: Teacher characteristics and student achievement. *The Elementary School Journal*, 99 (2), 101–128.

Authors' Note:

We thank Cristy Coughlin and Amy Johnston for their advice and comments on the design and analysis of this study.

Author Contact Information

Dr. Jean Stockard, Ph.D.
Stockard at the National Institute for Direct Instruction,
P.O. Box 11248
Eugene, Oregon, 97440
Phone 877-485-1972
fax: 541-683-7543
E-mail to jstockard@nifdi.org.

ADVERTISEMENT

ADVERTISING IN BAO Journals

If you wish to place an advertisement in any of our journals, you can do it by contacting us.

The prices for advertising in one issue are as follows:

1/4 Page: \$50.00 1/2 Page: \$100.00 Full Page: \$200.00

If you wish to run the same ad in multiple issues for the year, you are eligible for the following discount:

**1/4 Pg.: \$40 - per issue
1/2 Pg.: \$75 - per issue
Full Page: \$150.00-per issue**

An additional one time layout/composition fee of \$25.00 is applicable

In addition to placing your ad in the journal(s) of your choice, we will place your ad on our website's advertising section.

For more information, or place an ad, contact Halina Dziewolska by phone at (215) 462-6737 or e-mail at: halinadz@hotmail.com